

AFOSR 819

IMM-NYU 288
JUNE 1961



NEW YORK UNIVERSITY
INSTITUTE OF
MATHEMATICAL SCIENCES

A Machine Program For Theorem-Proving

MARTIN DAVIS, GEORGE LOGEMANN, and DONALD LOVELAND

REPRODUCTION IN WHOLE OR IN PART
IS PERMITTED FOR ANY PURPOSE
OF THE UNITED STATES GOVERNMENT.

PREPARED UNDER
CONTRACT NO. AF49(638)-777
MATHEMATICAL SCIENCES DIRECTORATE
AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

IMM-288
a.1

New York University
Institute of Mathematical Sciences
Mathematical Sciences Directorate
Air Force Office of Scientific Research
Washington 25, D.C.
AFOSR 819

A MACHINE PROGRAM FOR THEOREM-PROVING
Martin Davis, George Logemann, and Donald Loveland

June 8, 1961

Contract No. AF 49(638)-777

ABSTRACT: The programming of a proof procedure is discussed in connection with trial runs and possible improvements.

"Qualified requestors may obtain copies of this report from the ASTIA Document Service Center, Arlington Hall Station, Arlington 12, Virginia. Department of Defense contractors must be established for ASTIA services, or have their "need-to-know" certified by the cognizant military agency of their project or contract".

The research reported in this document has been sponsored by the Mathematical Sciences Directorate, Air Force Office of Scientific Research, under Contract No. AF 49(638)-777.

A Machine Program for Theorem-Proving
Martin Davis, George Logemann, and Donald Loveland.

In [1] is set forth an algorithm for proving theorems of quantification theory which is an improvement in certain respects over previously available algorithms such as that of [2]. The present paper deals with the programming of the algorithm of [1] for the New York University, Institute of Mathematical Sciences' IBM 704 computer, with some modifications in the algorithm suggested by this work, with the results obtained using the completed algorithm. Familiarity with [1] is assumed throughout.

1. Changes in the algorithm and programming techniques used.

The algorithm of [1] consists of two interlocking parts. The first part, called the QFL-Generator, generates (from the formula whose proof is being attempted) a growing propositional calculus formula in conjunctive normal form, the "quantifier-free lines." The second part, the Processor, tests, at regular stages in its "growth," the consistency of this propositional calculus formula. An inconsistent set of quantifier-free lines constitutes a proof of the original formula.

The algorithm of [1] used in testing for consistency proceeded by successive elimination of atomic formulas, first eliminating one-literal clauses (one-literal clause rule), and then atomic formulas all of whose occurrences were positive or all of whose occurrences were negative (affirmative-negative rule). Finally, the remaining atomic formulas were to have been eliminated by the rule:

III. Rule for Eliminating Atomic Formulas. Let the given formula F be put into the form $(A \vee p) \& (B \vee \bar{p}) \& R$ where A , B , and R are free of p . (This can be done simply by grouping

together the clauses containing p and "factoring out" occurrences of p to obtain A , grouping the clauses containing \bar{p} and "factoring out" \bar{p} to obtain B , and grouping the remaining clauses to obtain R .) Then F is inconsistent if and only if $(A \vee B) \& R$ is inconsistent.

After programming the algorithm using this form of Rule III, it was decided to replace it by the following rule:

III*. Splitting Rule. Let the given formula F be put in the form $(A \vee p) \& (B \vee \bar{p}) \& R$ where A, B , and R are free of p . Then F is inconsistent if and only if $A \& R$ and $B \& R$ are both inconsistent.

Justification of Rule III*. For¹ $p = 0$, $F = A \& R$; for $p = 1$, $F = B \& R$.

The forms of Rule III are interchangeable; although theoretically they are equivalent, in actual applications each has certain desirable features. We used Rule III* because of the fact that Rule III can easily increase the number and the lengths of the clauses in the expression rather quickly after several applications. This is prohibitive in a computer if one's fast access storage is limited. Also, it was observed that after performing Rule III, many duplicated and thus redundant clauses were present. Some success was had by causing the machine to systematically eliminate the redundancy; but the problem of total length increasing rapidly still remained when more complicated problems were attempted. Also use of Rule III can seldom yield new one-literal clauses, whereas use of Rule III* often will.

In programming Rule III*, we used auxiliary tape storage. The rest of the testing for consistency is carried out using only fast access storage. When the "Splitting Rule" is used one of the two formulas resulting is placed on tape. Tape memory records are organized in the cafeteria stack-of-plates scheme: the last record written is the first to be read.

In the program written for the IBM 704, the matrix and conjunction of quantifier-free lines are coded into cross-referenced associated (or linked) memory tables by the QFL-Generator and then analyzed by the Processor. In particular, the QFL-Generator is programmed to read in the matrix M in suitably coded Polish (i.e., "parenthesis-free") form. The conversion to a quantifier-free matrix in conjunctive normal form requires, of course, a certain amount of pencil work on the formula, which could have been done by the computer. In doing this, we departed from [1], by not using prenex normal form. The steps are :

(1) Write all truth-functional connectives in terms of \sim , $\&$, \vee .

(2) Move all \sim 's inward successively (using de Morgan laws) until they either are cancelled (with another \sim) or acting on an atomic formula.

(3) Now, replace existential quantifiers by function symbols (cf. [1], p. 205), drop universal quantifiers, and place in conjunctive normal form. A simple one-to-one assembler was written to perform the final translation of the matrix M into octal numbers.

It will be recalled that the generation of quantifier-free lines is accomplished by successive substitutions of "constants" for the variables in the matrix M. In the program the constants are represented by the successive positive integers.

For a matrix containing n individual variables, the n-tuples of positive integers are generated in a sequence of increasing norm such that all n-tuples with a given norm are in decreasing n-ary numerical order. Here we define the norm of $(j_1, \dots, j_n) = j_1 + \dots + j_n = ||j_i||$. Other norms could have been used.

For example, Gilmore [2] takes for $||j_i||$, the maximum of j_1, \dots, j_n . In [1] more complicated norm is indicated.

Substitutions of successive n-tuples into the matrix cause

new constants to appear in the matrix. The program numbers constants in their order of appearance. Thus, the constants are ordered by the program in a manner depending upon the input data. By rearranging the clauses of a formula a different order would in general be created. In some cases, whether or not the program could actually prove the validity of a given formula (without running out of fast access storage) depended on how one shuffled the punched-data deck before reading it into the assembler ! Thus, the variation in ordering of constants did affect by a factor of 10 (from 500 to 5000) the number of lines needed to prove the validity of :

$$(e)(Ed)(x)(y)[S(x,y,d) \longrightarrow T(x,y,e)]$$

$$\longrightarrow (e)(x)(Ed)(y)[S(x,y,d) \longrightarrow T(x,y,e)]$$

(This valid formula may be thought of as asserting that uniform continuity implies continuity if we set:

$$S(x,y,d) \longleftrightarrow |x - y| < d$$

$$T(x,y,e) \longleftrightarrow |f(x) - f(y)| < e.)$$

In storing the quantifier-free lines, two tables are used. The first, called the conjunction table, is a literal image of the quantifier-free lines in which one machine word corresponds to one literal, i.e. to p or $\sim p$ where p is an atomic formula. The lines in the second, or formula table are themselves heads of two chain lists giving the occurrences of p and $\sim p$ respectively in the conjunction table. In addition, included for formula p in the formula table are counts of the number of clauses in which p and $\sim p$ occur and total number of all literals in these clauses; the formula table is itself a two-way linked list. A third short list of those literals is kept in which are entered all formulas to which the one-literal clause and affirmative-negative rules apply; this is called the ready list.

If the program tries to enter p and $\sim p$ into the ready list, an inconsistency has been found; the machine stops.

The totality of the processing rules requires only two basic operations: a subroutine to delete the occurrences of a literal p or $\sim p$ from the quantifier-free lines, and a routine to eliminate from them all the clauses in which p or $\sim p$ occur.

We may observe that only the deletion program can create new one-literal clauses, and likewise applications of the affirmative-negative rule can come only from the elimination program.

The machine thus performs the one literal-clause and affirmative-negative rules as directed by the ready list until the ready list is empty. It is possible that the choice of p to be eliminated first is quite critical in determining the length of computation required to reach a conclusion: a program to choose p is used, but no tests were made to vary this segment of the program beyond a random selection, namely the first entry in the formula table. To perform Rule III*, one saves the appropriate tables with some added reference information in a tape record, then performs an elimination on $\sim p$ and a deletion on p . At a consequent discovery of consistency, one must generate more quantifier free lines; the QFL-generator is recalled. Otherwise, at finding an inconsistency, the machine must check to see if there are any records on the Rule III* tape: if none, the quantifier-free lines were inconsistent; otherwise, it reads in the last record.

If one uses Rule III, (which we did in an early version of our program) an entirely different code is needed. The problem is precisely that of mechanizing the application of the distributive law.

2. Results obtained in running the program.

At the time the programming of the algorithm was undertaken, we hoped that some mathematically meaningful and, perhaps non-trivial, theorems could be proved. The actual achievements in this direction were somewhat disappointing. However, the program

performed as well as expected on the simple predicate calculus formulas offered as fare for a previous proof procedure program.

(See Gilmore [1].) In particular, the well-formed formula

$$(Ex)(Ey)(z) \left\{ F(x,y) \rightarrow (F(y,z) \& F(z,z)) \& ((F(x,y) \& G(x,y)) \rightarrow (G(x,z) \& G(z,z))) \right\}$$

which was beyond the scope of Gilmore's program was proved in under two minutes with the present program. Gilmore's program was halted at the end of 7 "substitutions", (quantifier-free lines) after an elapsed period of about 21 minutes. It was necessary for the present program to generate approximately 60 quantifier-free lines before the inconsistency appeared.² Indeed, the "uniform continuity implies continuity" example mentioned above required over 500 quantifier-free lines to be generated and was shown to be valid in just over two minutes. This was accomplished by nearly filling the machine to capacity with generated quantifier-free lines (2000 lines in this case) before applying any of the rules of reduction .

Rather than describe further successes of the program, it will be instructive to consider in detail a theorem that the program was incapable of proving and to examine the cause for this. This particular example is one the authors originally had hoped the program could prove, an elementary group theory problem. In essence, it is to show that in a group a left inverse is also a right inverse.

It is, in fact, quite easy to follow the behavior of the

proof procedure on this particular example as it parallels the usual approach to the problem. The problem may be formulated as follows :

- Axioms:
1. $e \cdot x = x$
 2. $I(x) \cdot x = e$
 3. $(x \cdot y) \cdot z = w \Rightarrow x \cdot (y \cdot z) = w$
 4. $x \cdot (y \cdot z) = w \Rightarrow (x \cdot y) \cdot z = w$

Conclusion: $x \cdot I(x) = e$

The letter e is interpreted as the identity element and the function I as the inverse function. The associative law has been split into two clauses for convenience.

A proof is as follows :

1. $I(I(x)) \cdot I(x) = e$ by Axiom 2
2. $e \cdot x = x$ by Axiom 1
3. $I(x) \cdot x = e$ by Axiom 2
4. $I(I(x)) \cdot e = x$ by Axiom 3, taking $(I(I(x)), I(x), x)$ for (x, y, z)
5. $e \cdot I(x) = I(x)$ by Axiom 1
6. $I(I(x)) \cdot I(x) = e$ by Axiom 2
7. $I(I(x)) \cdot e = x$ step 4
8. $x \cdot I(x) = e$ by Axiom 4, taking $(I(I(x)), e, I(x))$ for (x, y, z)

Step 8. is the desired result.

To formalize this proof would require adjoining axioms of equality. To avoid this, one can introduce the predicate of

Journal of Management Studies, 1986, 23(7), 701-711

1. *Phragmites australis* (Cav.) Trin. ex Steud.

• • • • •

100

THE UNIVERSITY OF CHICAGO

... ..

$\sigma = \frac{1}{2}(\sigma_1 + \sigma_2)$

• Don't have a C rate to get into the club

1. 2000-2001

1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 2020 2021 2022 2023 2024 2025 2026 2027 2028 2029 2030 2031 2032 2033 2034 2035 2036 2037 2038 2039 2040 2041 2042 2043 2044 2045 2046 2047 2048 2049 2050 2051 2052 2053 2054 2055 2056 2057 2058 2059 2060 2061 2062 2063 2064 2065 2066 2067 2068 2069 2070 2071 2072 2073 2074 2075 2076 2077 2078 2079 2080 2081 2082 2083 2084 2085 2086 2087 2088 2089 2090 2091 2092 2093 2094 2095 2096 2097 2098 2099 2100 2101 2102 2103 2104 2105 2106 2107 2108 2109 2110 2111 2112 2113 2114 2115 2116 2117 2118 2119 2120 2121 2122 2123 2124 2125 2126 2127 2128 2129 2130 2131 2132 2133 2134 2135 2136 2137 2138 2139 2140 2141 2142 2143 2144 2145 2146 2147 2148 2149 2150 2151 2152 2153 2154 2155 2156 2157 2158 2159 2160 2161 2162 2163 2164 2165 2166 2167 2168 2169 2170 2171 2172 2173 2174 2175 2176 2177 2178 2179 2180 2181 2182 2183 2184 2185 2186 2187 2188 2189 2190 2191 2192 2193 2194 2195 2196 2197 2198 2199 2200 2201 2202 2203 2204 2205 2206 2207 2208 2209 2210 2211 2212 2213 2214 2215 2216 2217 2218 2219 2220 2221 2222 2223 2224 2225 2226 2227 2228 2229 2230 2231 2232 2233 2234 2235 2236 2237 2238 2239 2240 2241 2242 2243 2244 2245 2246 2247 2248 2249 2250 2251 2252 2253 2254 2255 2256 2257 2258 2259 2260 2261 2262 2263 2264 2265 2266 2267 2268 2269 2270 2271 2272 2273 2274 2275 2276 2277 2278 2279 2280 2281 2282 2283 2284 2285 2286 2287 2288 2289 2290 2291 2292 2293 2294 2295 2296 2297 2298 2299 2300 2301 2302 2303 2304 2305 2306 2307 2308 2309 2310 2311 2312 2313 2314 2315 2316 2317 2318 2319 2320 2321 2322 2323 2324 2325 2326 2327 2328 2329 2330 2331 2332 2333 2334 2335 2336 2337 2338 2339 2340 2341 2342 2343 2344 2345 2346 2347 2348 2349 2350 2351 2352 2353 2354 2355 2356 2357 2358 2359 2360 2361 2362 2363 2364 2365 2366 2367 2368 2369 2370 2371 2372 2373 2374 2375 2376 2377 2378 2379 2380 2381 2382 2383 2384 2385 2386 2387 2388 2389 2390 2391 2392 2393 2394 2395 2396 2397 2398 2399 2400 2401 2402 2403 2404 2405 2406 2407 2408 2409 2410 2411 2412 2413 2414 2415 2416 2417 2418 2419 2420 2421 2422 2423 2424 2425 2426 2427 2428 2429 2430 2431 2432 2433 2434 2435 2436 2437 2438 2439 2440 2441 2442 2443 2444 2445 2446 2447 2448 2449 2450 2451 2452 2453 2454 2455 2456 2457 2458 2459 2460 2461 2462 2463 2464 2465 2466 2467 2468 2469 2470 2471 2472 2473 2474 2475 2476 2477 2478 2479 2480 2481 2482 2483 2484 2485 2486 2487 2488 2489 2490 2491 2492 2493 2494 2495 2496 2497 2498 2499 2500 2501 2502 2503 2504 2505 2506 2507 2508 2509 2510 2511 2512 2513 2514 2515 2516 2517 2518 2519 2520 2521 2522 2523 2524 2525 2526 2527 2528 2529 2530 2531 2532 2533 2534 2535 2536 2537 2538 2539 2540 2541 2542 2543 2544 2545 2546 2547 2548 2549 2550 2551 2552 2553 2554 2555 2556 2557 2558 2559 2560 2561 2562 2563 2564 2565 2566 2567 2568 2569 2570 2571 2572 2573 2574 2575 2576 2577 2578 2579 2580 2581 2582 2583 2584 2585 2586 2587 2588 2589 2590 2591 2592 2593 2594 2595 2596 2597 2598 2599 2600 2601 2602 2603 2604 2605 2606 2607 2608 2609 2610 2611 2612 2613 2614 2615 2616 2617 2618 2619 2620 2621 2622 2623 2624 2625 2626 2627 2628 2629 2630 2631 2632 2633 2634 2635 2636 2637 2638 2639 2640 2641 2642 2643 2644 2645 2646 2647 2648 2649 2650 2651 2652 2653 2654 2655 2656 2657 2658 2659 2660 2661 2662 2663 2664 2665 2666 2667 2668 2669 2670 2671 2672 2673 2674 2675 2676 2677 2678 2679 2680 2681 2682 2683 2684 2685 2686 2687 2688 2689 2690 2691 2692 2693 2694 2695 2696 2697 2698 2699 2700 2701 2702 2703 2704 2705 2706 2707 2708 2709 2710 2711 2712 2713 2714 2715 2716 2717 2718 2719 2720 2721 2722 2723 2724 2725 2726 2727 2728 2729 2730 2731 2732 2733 2734 2735 2736 2737 2738 2739 2740 2741 2742 2743 2744 2745 2746 2747 2748 2749 2750 2751 2752 2753 2754 2755 2756 2757 2758 2759 2760 2761 2762 2763 2764 2765 2766 2767 2768 2769 2770 2771 2772 2773 2774 2775 2776 2777 2778 2779 2780 2781 2782 2783 2784 2785 2786 2787 2788 2789 2790 2791 2792 2793 2794 2795 2796 2797 2798 2799 2800 2801 2802 2803 2804 2805 2806 2807 2808

THE UNIVERSITY OF CHICAGO

1997, 1998, 1999, 2000, 2001, 2002, 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2010, 2011, 2012, 2013, 2014, 2015, 2016, 2017, 2018, 2019, 2020, 2021, 2022, 2023, 2024, 2025, 2026, 2027, 2028, 2029, 2030, 2031, 2032, 2033, 2034, 2035, 2036, 2037, 2038, 2039, 2040, 2041, 2042, 2043, 2044, 2045, 2046, 2047, 2048, 2049, 2050, 2051, 2052, 2053, 2054, 2055, 2056, 2057, 2058, 2059, 2060, 2061, 2062, 2063, 2064, 2065, 2066, 2067, 2068, 2069, 2070, 2071, 2072, 2073, 2074, 2075, 2076, 2077, 2078, 2079, 2080, 2081, 2082, 2083, 2084, 2085, 2086, 2087, 2088, 2089, 2090, 2091, 2092, 2093, 2094, 2095, 2096, 2097, 2098, 2099, 2100, 2101, 2102, 2103, 2104, 2105, 2106, 2107, 2108, 2109, 2110, 2111, 2112, 2113, 2114, 2115, 2116, 2117, 2118, 2119, 2120, 2121, 2122, 2123, 2124, 2125, 2126, 2127, 2128, 2129, 2130, 2131, 2132, 2133, 2134, 2135, 2136, 2137, 2138, 2139, 2140, 2141, 2142, 2143, 2144, 2145, 2146, 2147, 2148, 2149, 2150, 2151, 2152, 2153, 2154, 2155, 2156, 2157, 2158, 2159, 2160, 2161, 2162, 2163, 2164, 2165, 2166, 2167, 2168, 2169, 2170, 2171, 2172, 2173, 2174, 2175, 2176, 2177, 2178, 2179, 2180, 2181, 2182, 2183, 2184, 2185, 2186, 2187, 2188, 2189, 2190, 2191, 2192, 2193, 2194, 2195, 2196, 2197, 2198, 2199, 2200, 2201, 2202, 2203, 2204, 2205, 2206, 2207, 2208, 2209, 2210, 2211, 2212, 2213, 2214, 2215, 2216, 2217, 2218, 2219, 2220, 2221, 2222, 2223, 2224, 2225, 2226, 2227, 2228, 2229, 2230, 2231, 2232, 2233, 2234, 2235, 2236, 2237, 2238, 2239, 2240, 2241, 2242, 2243, 2244, 2245, 2246, 2247, 2248, 2249, 2250, 2251, 2252, 2253, 2254, 2255, 2256, 2257, 2258, 2259, 2260, 2261, 2262, 2263, 2264, 2265, 2266, 2267, 2268, 2269, 2270, 2271, 2272, 2273, 2274, 2275, 2276, 2277, 2278, 2279, 2280, 2281, 2282, 2283, 2284, 2285, 2286, 2287, 2288, 2289, 2290, 2291, 2292, 2293, 2294, 2295, 2296, 2297, 2298, 2299, 2300, 2301, 2302, 2303, 2304, 2305, 2306, 2307, 2308, 2309, 2310, 2311, 2312, 2313, 2314, 2315, 2316, 2317, 2318, 2319, 2320, 2321, 2322, 2323, 2324, 2325, 2326, 2327, 2328, 2329, 2330, 2331, 2332, 2333, 2334, 2335, 2336, 2337, 2338, 2339, 2340, 2341, 2342, 2343, 2344, 2345, 2346, 2347, 2348, 2349, 2350, 2351, 2352, 2353, 2354, 2355, 2356, 2357, 2358, 2359, 2360, 2361, 2362, 2363, 2364, 2365, 2366, 2367, 2368, 2369, 2370, 2371, 2372, 2373, 2374, 2375, 2376, 2377, 2378, 2379, 2380, 2381, 2382, 2383, 2384, 2385, 2386, 2387, 2388, 2389, 2390, 2391, 2392, 2393, 2394, 2395, 2396, 2397, 2398, 2399, 2400, 2401, 2402, 2403, 2404, 2405, 2406, 2407, 2408, 2409, 2410, 2411, 2412, 2413, 2414, 2415, 2416, 2417, 2418, 2419, 2420, 2421, 2422, 2423, 2424, 2425, 2426, 2427, 2428, 2429, 2430, 2431, 2432, 2433, 2434, 2435, 2436, 2437, 2438, 2439, 2440, 2441, 2442, 2443, 2444, 2445, 2446, 2447, 2448, 2449, 2450, 2451, 2452, 2453, 2454, 2455, 2456, 2457, 2458, 2459, 2460, 2461, 2462, 2463, 2464, 2465, 2466, 2467, 2468, 2469, 2470, 2471, 2472, 2473, 2474, 2475, 2476, 2477, 2478, 2479, 2480, 2481, 2482, 2483, 2484, 2485, 2486, 2487, 2488, 2489, 2490, 2491, 2492, 2493, 2494, 2495, 2496, 2497, 2498, 2499, 2500, 2501, 2502, 2503, 2504, 2505, 2506, 2507, 2508, 2509, 2510, 2511, 2512, 2513, 2514, 2515, 2516, 2517, 2518, 2519, 2520, 2521, 2522, 2523, 2524, 2525, 2526, 2527, 2528, 2529, 2530, 2531, 2532, 2533, 2534, 2535, 2536, 2537, 2538, 2539, 2540, 2541, 2542, 2543, 2544, 2545, 2546, 2547, 2548, 2549, 2550, 2551, 2552, 2553, 2554, 2555, 2556, 2557, 2558, 2559, 2560, 2561, 2562, 2563, 2564, 2565, 2566, 2567, 2568, 2569, 2570, 2571, 2572, 2573, 2574, 2575, 2576, 2577, 2578, 2579, 2580, 2581, 2582, 2583, 2584, 2585, 2586, 2587, 2588, 2589, 2590, 2591, 2592, 2593, 2594, 2595, 2596, 2597, 2598, 2599, 2600, 2601, 2602, 2603, 2604, 2605, 2606, 2607, 2608, 2609, 2610, 2611, 2612, 2613, 2614, 2615, 2616, 2617, 2618, 2619, 2620, 2621, 2622, 2623, 2624, 2625, 2626, 2627, 2628, 2629, 2630, 2631, 2632, 2633, 2634, 2635, 2636, 2637, 2638, 2639, 2640, 2641, 2642, 2643, 2644, 2645, 2646, 2647, 2648, 2649, 2650, 2651, 2652, 2653, 2654, 2655, 2656, 2657, 2658, 2659, 2660, 2661, 2662, 2663, 2664, 2665, 2666, 2667, 2668, 2669, 2670, 2671, 2672, 2673, 2674, 2675, 2676, 2677, 2678, 26

... ..

[illegible]

1950 年 10 月 1 日 至 1951 年 9 月 30 日 共 365 天

three arguments $P(x,y,z)$, interpreted as $x \cdot y = z$. The theorem reformulated becomes

- Axioms:
1. $P(e,x,x)$
 2. $P(I(x),x,e)$
 3. $\sim P(x,y,u) \vee \sim P(u,z,w) \vee \sim P(y,z,v) \vee P(x,v,w)$
 4. $\sim P(y,z,v) \vee \sim P(x,v,w) \vee \sim P(x,y,u) \vee P(u,z,w)$

Conclusion : $P(x,I(x),e)$.

The theorem to be proved valid is the implication of the conjunction of the four axioms with the conclusion, the universal quantifiers appearing outside the matrix.

To complete the preparation of the well-formed formula for encoding for the assembler, it is necessary to negate the conclusion. (cf. [1], p. 204).

The single existential quantifier has no dependence on the universal quantifiers, hence leads to the constant function s when this existential quantifier is replaced by a function symbol. (cf. [1], p. 205.)

The conclusion then becomes

$$\sim P(s, I(s), e) .$$

The conjunction of this with the 4 axioms gives the desired form.

As seen from the proof previously noted the quantifier-free clauses needed to produce the inconsistency are

1. $P(I(I(s)), I(s), e)$
2. $P(e, s, s)$
3. $P(I(s), s, e)$

... ..

... ..

... ..

... ..

... ..

... ..

... ..

... ..

... ..

... ..

... ..

... ..

... ..

... ..

... ..

... ..

... ..

... ..

... ..

... ..

... ..

... ..

... ..

... ..

... ..

4. $\sim P(I(I(s)), I(s), e) \vee \sim P(e, s, s) \vee \sim P(I(s), s, e) \vee P(I(I(s)), e, s)$
5. $P(e, I(s), I(s))$
6. $\sim P(e, I(s), I(s)) \vee \sim P(I(I(s)), I(s), e) \vee \sim P(I(I(s)), e, s) \vee P(s, I(s), e)$
7. $\sim P(s, I(s), e)$

(It is quite clear in this case that successive applications of the one-literal clause rule reduce this set to:

$$P(s, I(s), e) \& \sim P(s, I(s), e).)$$

The question to be considered is: how many quantifier-free lines must be generated by the present program to realize these required lines? The constants are generated in the following order:

1. e
2. s
3. $I(s)$
4. $I(e)$
5. $I(I(s))$
- etc.

(The constants are identified directly with their index e.g. the 6-tuple (1,1,1,1,1,1) represents (e,e,e,e,e,e). As this is the first substitution, the program assigns in order, reading the well-formed formula backwards and from the inside out for nested functions; $e, s, I(s), I(e), I(I(s))$. The $I(I(s))$ appears when x is assigned $I(s)$, no new entries occurring until this time. Note that there are 6 free variables (u, v, w, x, y, z)

1. The first part of the document is a list of names and addresses of the members of the committee.

(1) Mr. J. H. Smith, 123 Main St., New York, N.Y.

(2) Mr. J. H. Smith, 123 Main St., New York, N.Y.

(3) Mr. J. H. Smith, 123 Main St., New York, N.Y.

(4) Mr. J. H. Smith, 123 Main St., New York, N.Y.

2. The second part of the document is a list of names and addresses of the members of the committee.

(1) Mr. J. H. Smith, 123 Main St., New York, N.Y.

(2) Mr. J. H. Smith, 123 Main St., New York, N.Y.

3. The third part of the document is a list of names and addresses of the members of the committee.

(1) Mr. J. H. Smith, 123 Main St., New York, N.Y.

(2) Mr. J. H. Smith, 123 Main St., New York, N.Y.

(3) Mr. J. H. Smith, 123 Main St., New York, N.Y.

(4) Mr. J. H. Smith, 123 Main St., New York, N.Y.

(5) Mr. J. H. Smith, 123 Main St., New York, N.Y.

(6) Mr. J. H. Smith, 123 Main St., New York, N.Y.

(7) Mr. J. H. Smith, 123 Main St., New York, N.Y.

(8) Mr. J. H. Smith, 123 Main St., New York, N.Y.

(9) Mr. J. H. Smith, 123 Main St., New York, N.Y.

4. The fourth part of the document is a list of names and addresses of the members of the committee.

(1) Mr. J. H. Smith, 123 Main St., New York, N.Y.

(2) Mr. J. H. Smith, 123 Main St., New York, N.Y.

(3) Mr. J. H. Smith, 123 Main St., New York, N.Y.

(4) Mr. J. H. Smith, 123 Main St., New York, N.Y.

(5) Mr. J. H. Smith, 123 Main St., New York, N.Y.

(6) Mr. J. H. Smith, 123 Main St., New York, N.Y.

(7) Mr. J. H. Smith, 123 Main St., New York, N.Y.

in the matrix).

The program generates the needed n -tuples by producing all possible n -tuples of integers whose sum N of entries is fixed, $N=n, n+1, \dots$. Thus it is only necessary to consider the n -tuple which has the maximum sum of entries. In this case, the substitution $u = s, v = I(s), w = e, x = I(I(s)), y = e, z = I(s)$ (required for axiom 4 to produce the clause 6 in the "proof" above in a quantifier-free line) gives the n -tuple with maximum sum. The n -tuple is seen to be $(2,3,1,5,1,3)$, the sum equals 15. The combinatorial expression $\binom{N}{n}$ gives the total number of n -tuples of positive integers whose sum is less than or equal to N .³

It is seen that to prove this theorem at least $\binom{14}{6} = 3003$ lines must be generated and that the inconsistency will be found on or before $\binom{15}{6} = 5005$ lines have been generated.⁴

The present program generated approximately 1300 quantifier free lines. This number of quantifier-free lines was accomplished holding all major tables simultaneously in core memory, limited to 32,768 "words". (This was done to insure a reasonable time factor for any problem within possible scope of the program.

For this reason also, the entire program was coded in SAP with many time-saving devices employed.)

The authors feel that a reprogramming to make use of tape storage of tables might realize a factor of 4 for the total number of quantifier-free lines attainable before running time became prohibitive. This would be just sufficient for this problem. That realizing this extra capacity is really uninter-

The second condition is that the function f is continuous on $[a, b]$.

Let us now consider the case where f is not continuous on $[a, b]$.

Suppose that f is continuous on $[a, c]$ and $[c, b]$, but not at c .

Let $F(x)$ be the function defined by

$$F(x) = \begin{cases} f(t) dt & \text{for } a \leq t \leq c \\ f(t) dt & \text{for } c \leq t \leq b \end{cases}$$

Then $F(x)$ is continuous on $[a, b]$ and $F'(x) = f(x)$ for all x in $[a, b]$.

Thus, the function $F(x)$ is an antiderivative of $f(x)$ on $[a, b]$.

It is easy to see that if f is continuous on $[a, b]$, then

$$\int_a^b f(x) dx = F(b) - F(a)$$

where $F(x)$ is any antiderivative of $f(x)$ on $[a, b]$.

It is also true that if f is not continuous on $[a, b]$, then

$$\int_a^b f(x) dx \neq F(b) - F(a)$$

where $F(x)$ is any antiderivative of $f(x)$ on $[a, b]$.

The reason for this is that if f is not continuous on $[a, b]$, then

$$\int_a^b f(x) dx \neq \int_a^c f(x) dx + \int_c^b f(x) dx$$

where c is a point in $[a, b]$ where f is not continuous.

Thus, the function $F(x)$ is not an antiderivative of $f(x)$ on $[a, b]$.

Therefore, the function $F(x)$ is not an antiderivative of $f(x)$ on $[a, b]$.

Thus, the function $F(x)$ is not an antiderivative of $f(x)$ on $[a, b]$.

Thus, the function $F(x)$ is not an antiderivative of $f(x)$ on $[a, b]$.

Thus, the function $F(x)$ is not an antiderivative of $f(x)$ on $[a, b]$.

Thus, the function $F(x)$ is not an antiderivative of $f(x)$ on $[a, b]$.

Thus, the function $F(x)$ is not an antiderivative of $f(x)$ on $[a, b]$.

Thus, the function $F(x)$ is not an antiderivative of $f(x)$ on $[a, b]$.

Thus, the function $F(x)$ is not an antiderivative of $f(x)$ on $[a, b]$.

esting is seen by noting that if the conclusion was placed before the axioms, altering the validity of the matrix not at all, the element $I(e)$ would be generated before $I(s)$ and the needed n -tuple would sum to 16. Then $\binom{16}{6} = 8008$ becomes the upper bound, beyond the capacity of the projected program. Other formulations of the same problem result in quite unapproachable figures for the number of quantifier-free lines needed. (For another example illustrating the same situation, see Prawitz [3]).

The existing program allows one to think of working with a capacity of 1000 or 2000 quantifier-free lines instead of a capacity of 10 or 20, the previous limit. The time required to generate additional quantifier-free lines is independent of the number of quantifier-free lines already existing. Against this linear growth of number of quantifier-free lines generated, there is, in a meaningful sense, an extreme non-linear growth in the number of quantifier-free lines to be considered with increasingly more "difficult" problems. This is true of simple enumeration schemes of the nature considered here. It seems that the most fruitful future results will come from reducing the number of quantifier-free lines that need be considered, by excluding, in some sense, "irrelevant" quantifier-free lines. Some investigation in this area has already been done (see Prawitz [3]).

Footnotes.

¹ As in [1], 1 stands for "truth", and 0 for "falsehood".

² In [1], a hand-calculation of this example using the present scheme showed inconsistency at 25 quantifier-free lines. The discrepancy is due to a different rule for generation of constants.

³ To see this, consider a sequence of $N + 1$ ones. Flag n of these. The flag is to be interpreted "sum all 1's, including the flagged 1, to the next flag and consider this sum as an entry in the n -tuple". Placing an unflagged 1 on the extreme left, leaving it fixed, consider the possible permutations of all other symbols. The different sequences total $\binom{N}{n}$ and, regarding the set of 1's starting with the last flagged 1 as overflow, this is seen to represent precisely the desired n -tuples.

⁴ If the rule for generating n -tuples had been, for each m , to generate all n -tuples possible such that each entry assumes a positive integral value $\leq m$, it is clear that at least $4^6 = 4096$ quantifier-free lines would be needed and $5^6 = 15625$ lines would suffice to guarantee a solution. If no more information were available, one sees an intuitive advantage, in this case, for using the previous method. In general, the authors see no preference for either method, in contrast to some previous suggestions that the latter method seemed intuitively better.

1. The first part of the paper is devoted to the study of the

properties of the function $f(x)$ defined by the equation

$f(x) = \frac{1}{2} (f(x-1) + f(x+1))$.

It is shown that the function $f(x)$ is a linear function of x .

2. In the second part of the paper, we consider the function

$g(x) = \frac{1}{2} (g(x-1) + g(x+1))$.

It is shown that the function $g(x)$ is a linear function of x .

3. In the third part of the paper, we consider the function

$h(x) = \frac{1}{2} (h(x-1) + h(x+1))$.

It is shown that the function $h(x)$ is a linear function of x .

4. In the fourth part of the paper, we consider the function

$k(x) = \frac{1}{2} (k(x-1) + k(x+1))$.

It is shown that the function $k(x)$ is a linear function of x .

5. In the fifth part of the paper, we consider the function

$l(x) = \frac{1}{2} (l(x-1) + l(x+1))$.

It is shown that the function $l(x)$ is a linear function of x .

6. In the sixth part of the paper, we consider the function

$m(x) = \frac{1}{2} (m(x-1) + m(x+1))$.

It is shown that the function $m(x)$ is a linear function of x .

7. In the seventh part of the paper, we consider the function

$n(x) = \frac{1}{2} (n(x-1) + n(x+1))$.

REFERENCES

- [1] Martin Davis and Hilary Putnam, A Computing Procedure for Quantification Theory. J. ACM Vol. 7, No. 3, (1960), 201-215.
- [2] P.C. Gilmore, A Proof Method for Quantification Theory. IBM J. Research and Development, Vol. 4 (1960) 28-35.
- [3] Dag Prawitz, An Improved Proof Procedure. Theoria, Vol. XXVI, 2; (1960), 102-139.

DISTRIBUTION LIST AIR FORCE OFFICE OF SCIENTIFIC RESEARCH
MATHEMATICAL SCIENCES DIRECTORATE
(ONE COPY UNLESS OTHERWISE NOTED)

ALABAMA

Commander
Army Rocket & Guided Missile Agency
ATTN: ORDXR-OTL
Redstone Arsenal, Alabama

BELGIUM

Commander (3)
European Office, ARDC
47 Rue Cantersteen
Brussels, Belgium

CALIFORNIA

Applied Mathematics & Statistics
Laboratory
Stanford University
Stanford, California

Department of Mathematics
University of California
Berkeley, California

Commander
Air Force Flight Test Center
Attn: Technical Library
Edwards Air Force Base, California

The Rand Corporation (2)
Technical Library
1700 Main Street
Santa Monica, California

Commander
1st Missile Division
ATTN: Operations Analysis Office
Vandenberg Air Force Base,
California

CONNECTICUT

Department of Mathematics
Yale University
New Haven, Connecticut

DISTRICT OF COLUMBIA

Office of Naval Research (2)
Department of the Navy
ATTN: Code 432
Washington 25, D.C.

Director
Department of Commerce
Office of Technical Services
Washington 25, D.C.

Administrator (6)
National Aeronautics &
Space Administration
ATTN: Documents Library
1520 H Street, N. W.
Washington 25, D. C.

Library
National Bureau of Standards
Washington 25, D. C.

Data Processing Systems
Division
National Bureau of Standards
ATTN: Mr. Russell A. Kirsch
Washington 25, D. C.

Applied Mathematics Division
National Bureau of Standards
Washington 25, D. C.

Headquarters, USAF
Assistant for Operations
Analysis
Deputy Chief of Staff,
Operations, AF00A
Washington 25, D. C.

Commander (2)
Air Force Office of Scientific
Research
ATTN: SRM
Washington 25, D. C.

Director
U. S. Naval Research Laboratory
ATTN: Library
Washington 25, D. C.

Commander, AFRD (2)
ATTN: Technical Library
Washington 25, D. C.

National Science Foundation
Program Director for Mathematical
Sciences
Washington 25, D. C.

Canadian Joint Staff
ATTN: DRB/DSIS
2450 Massachusetts Avenue, N. W.
Washington 25, D. C.

ILLINOIS

Department of Mathematics
Northwestern University
Evanston, Illinois

Laboratories for Applied Sciences
University of Chicago
Museum of Science and Industry
ATTN: Library, W-305
Chicago 37, Illinois

Department of Mathematics
University of Chicago
Chicago 37, Illinois

Department of Mathematics
University of Illinois
Urbana, Illinois

INDIANA

Department of Mathematics
Purdue University
Lafayette, Indiana

MARYLAND

Institute for Fluid Dynamics
and Applied Mathematics
University of Maryland
College Park, Maryland

Mathematics and Physics Library
The Johns Hopkins University
Baltimore, Maryland

Director
National Security Agency
ATTN: Dr. H. H. Campaign
Fort George G. Meade, Maryland

MASSACHUSETTS

Department of Mathematics
Harvard University
Cambridge 38, Massachusetts

Department of Mathematics
Massachusetts Institute of
Technology
Cambridge 38, Massachusetts

Commander
Detachment 2, AFRD
ATTN: Technical Library
L. G. Hanscom Field
Bedford, Massachusetts

MICHIGAN

Department of Mathematics
Wayne State University
Detroit 1, Michigan

MINNESOTA

Department of Mathematics
Folwell Hall
University of Minnesota
Minneapolis, Minnesota

Department of Mathematics
Institute of Technology
Engineering Building
University of Minnesota
Minneapolis, Minnesota

MISSOURI

Department of Mathematics
Washington University
St. Louis 8, Missouri

Department of Mathematics
University of Missouri
Columbia, Missouri

1. The first part of the document
describes the general situation
of the country.

2. The second part of the document
describes the situation in the
different regions.

3. The third part of the document
describes the situation in the
different regions.

4. The fourth part of the document
describes the situation in the
different regions.

5. The fifth part of the document
describes the situation in the
different regions.

6. The sixth part of the document
describes the situation in the
different regions.

7. The seventh part of the document
describes the situation in the
different regions.

8. The eighth part of the document
describes the situation in the
different regions.

9. The ninth part of the document
describes the situation in the
different regions.

10. The tenth part of the document
describes the situation in the
different regions.

1. The first part of the document
describes the general situation
of the country.

2. The second part of the document
describes the situation in the
different regions.

3. The third part of the document
describes the situation in the
different regions.

4. The fourth part of the document
describes the situation in the
different regions.

5. The fifth part of the document
describes the situation in the
different regions.

6. The sixth part of the document
describes the situation in the
different regions.

7. The seventh part of the document
describes the situation in the
different regions.

8. The eighth part of the document
describes the situation in the
different regions.

9. The ninth part of the document
describes the situation in the
different regions.

10. The tenth part of the document
describes the situation in the
different regions.

NEBRASKA

Commander
Strategic Air Command
ATTN: Operations Analysis
Offutt Air Force Base
Omaha, Nebraska

NEW JERSEY

The James Forrestal Research
Center Library
Princeton University
Princeton, New Jersey

Library
Institute for Advanced Study
Princeton, New Jersey

Department of Mathematics
Fine Hall
Princeton University
Princeton, New Jersey

Commanding General
Signal Corps Engineering Laboratory
ATTN: SIGFM/EL-RPO
Ft. Monmouth, New Jersey

NEW MEXICO

Commander
Air Force Missile Development
Center
ATTN: Technical Library, HDOI
Holloman Air Force Base, N.M.

Commander
Air Force Special Weapons Center
ATTN: Technical Library, SMOI
Kirtland Air Force Base
Albuquerque, New Mexico

NEW YORK

Professor J. Wolfowitz
Mathematics Department
White Hall
Cornell University
Ithaca, New York

Department of Mathematics
Syracuse University
Syracuse, New York

Institute for Mathematical
Sciences
New York University
ATTN: Professor M. Kline
25 Waverly Place
New York 3, New York

Institute for Aeronautical
Sciences
ATTN: Librarian
2 East 64th Street
New York 16, New York

NORTH CAROLINA

Department of Mathematics
University of North Carolina
Chapel Hill, North Carolina

Department of Statistics
University of North Carolina
Chapel Hill, North Carolina

Office of Ordnance Research
Box CM (2)
Duke Station
Durham, North Carolina

Department of Mathematics
Duke University
Duke Station
Durham, North Carolina

OHIO

P. O. Box AA
Wright-Patterson Air Force
Base
Ohio

Commander
Wright Air Development Division
ATTN: WCOSI
Wright-Patterson Air Force
Base
Ohio

Commander
Aeronautical Research
Laboratories
ATTN: Technical Library
Wright-Patterson Air Force
Base
Ohio

1. The first part of the report
describes the general situation
of the country and the
state of the economy.

2. The second part of the report
describes the situation in the
different regions of the country.

3. The third part of the report
describes the situation in the
different sectors of the economy.

4. The fourth part of the report
describes the situation in the
different social groups.

(S) 5. The fifth part of the report
describes the situation in the
different political groups.

6. The sixth part of the report
describes the situation in the
different cultural groups.

7. The seventh part of the report
describes the situation in the
different religious groups.

8. The eighth part of the report
describes the situation in the
different ethnic groups.

9. The ninth part of the report
describes the situation in the
different linguistic groups.

10. The tenth part of the report
describes the situation in the
different geographical groups.

1. The first part of the report
describes the general situation
of the country and the
state of the economy.

2. The second part of the report
describes the situation in the
different regions of the country.

3. The third part of the report
describes the situation in the
different sectors of the economy.

4. The fourth part of the report
describes the situation in the
different social groups.

5. The fifth part of the report
describes the situation in the
different political groups.

6. The sixth part of the report
describes the situation in the
different cultural groups.

7. The seventh part of the report
describes the situation in the
different religious groups.

8. The eighth part of the report
describes the situation in the
different ethnic groups.

9. The ninth part of the report
describes the situation in the
different linguistic groups.

10. The tenth part of the report
describes the situation in the
different geographical groups.

USAF Institute of Technology (2)
Library
ATTN: MCLI-ITLIB
Building 125, Area B
Wright-Patterson Air Force Base
Ohio

Mathematics Research Center,
U. S. Army
ATTN: R. E. Langer
University of Wisconsin
Madison, Wisconsin

PENNSYLVANIA

Department of Mathematics
Carnegie Institute of Technology
Pittsburgh, Pennsylvania

Department of Mathematics
University of Pennsylvania
Philadelphia, Pennsylvania

TENNESSEE

AEDC Library
ARO, Inc.
Arnold AF Station, Tennessee

U. S. Atomic Energy Commission
Technical Information Service
Extension
P. O. Box 62
Oak Ridge, Tennessee

TEXAS

Applied Mechanics Reviews (2)
Southwest Research Institute
8500 Culebra Road
San Antonio 6, Texas

Department of Mathematics
Rice Institute
Houston, Texas

VIRGINIA

Armed Services Technical (2)
Information Agency
ATTN : TIPDR
Arlington Hall Station
Arlington 12, Virginia

WISCONSIN

Department of Mathematics
University of Wisconsin
Madison, Wisconsin

4 Washington Place, New York 3, N. Y.

Date Due

[illegible]

PRINTED	IN U. S. A.
---------	-------------

NYU
IMM-288

c.1

Davis

A machine program for
theorem-proving

NYU
IMM-288

c.1

Davis

AUTHOR

A machine program for

TITLE

theorem-proving

DATE DUE

BORROWER'S NAME

**N. Y. U. Institute of
Mathematical Sciences**

4 Washington Place
New York 3, N. Y.

